A LRF and stereovision based data association method for objects tracking

Pawel Kmiotek, Cyril Meurie, Yassine Ruichek, Frederick Zann

A LRF and stereovision based data association method for objects tracking

Pawel Kmiotek and Cyril Meurie and Yassine Ruichek and Frederick Zann
Systems and Transportation Laboratory, University of Technology of Belfort-Montbeliard
13 rue Ernest-Thierry Mieg, 90010 Belfort Cedex, France
{pawel.kmiotek, cyril.meurie, yassine.ruichek, frederick.zann}@utbm.fr

***Abstract*—This paper presents a fusion method for objects tracking using laser sensory data and stereovision. Based on the Extended Kalman filter, the tracking uses an Oriented Bounding Box (OBB) representation for tracked objects. The representation model takes into account an Inter-Rays (IR) uncertainty concept, which is related to the fact that the laser raw data points representing the extremities of an extracted OBB do not coincide with the real objects extremities. To improve the objects state estimation, the tracking process integrates a Fixed Size (FS) assumption. The FS assumption allows to exploit the most precise object's size estimation, memorised during the tracking. To achieve data association, a threshold based laser points clustering provides satisfying results. However, there are many cases where, without additional information, it is impossible to cluster laser raw data points correctly. To discard clustering ambiguities, a fusion method combining laser sensory data and stereovision information is proposed. The stereovision information is extracted only within regions of interest, defined from laser points. The fusion method takes place in the early stage of the measurement extraction from laser raw data points. The proposed approach is tested and evaluated to demonstrate its reliability.**

***Index Terms*—data fusion, data association, laser scanner, stereovision, object tracking, intelligent vehicle**

I. Introduction

In the last decade, many research programs have been launched to study the concept of intelligent vehicles and their integration in the city of the future. The goal is to develop intelligent transportation systems (autonomous driving, platoon, etc.) having the ability to navigate autonomously in various urban environments. To reach the objective, the first primary task is to develop a perception system for detecting, localising and tracking objects in this type of environment. In this framework, the Systems and Transportation Laboratory of the University of Technology of Belfort-Montbeliard (France) develops a research program based on an experimental platform consisting of an electrical vehicle with an automatic control, equipped with several sensors and communication interfaces. In this paper, the emphasis is put on tracking of compact dynamic objects using laser sensory data and stereovision information.

Representation of dynamic objects is crucial for tracking and trajectory planning. Oriented Bounding Box (OBB) is chosen as an optimal model to represent tracked objects in urban environment [1][2]. Indeed, the OBB representation provides a good approximation of the area occupied by the tracked objects and a good data compression ratio. To increase the precision of the OBB extraction, an Inter-Rays (IR) uncertainty is used to take into account the fact that the laser raw data points representing the extremities of an extracted OBB do not coincide with the real objects extremities.

Tracking objects depends directly on data association, which starts by data clustering. The clustering leads to one of the three following situations: new object to track, single object tracking and multiple objects tracking. LRF data based tracking provides generally precise and fast objects sate estimation, using threshold clustering [3][4][5]. However, for many cases, laser sensory data based data association cannot lead to a correct clustering, and, hence, the tracking fails. To take into account clustering ambiguous cases, the authors propose to enrich laser sensory data using stereovision information.

The proposed tracking system is based on the Extended Kalman Filter (EKF) with Discrete White Noise Acceleration Model (DWNA) [6]. In order to increase its reliability, the tracking process integrates a Fixed Size (FS) assumption, which is introduced to exploit the most precise object's size estimation, memorised during the tracking.

The paper is organised as follows. Section II presents the object representation model. Section III describes the laser data association method, combined with stereovision information. In section IV, the tracking procedure is explained. Before concluding, experimental results are presented in section V to demonstrate the effectiveness and the reliability of the proposed approach.

II. Object representation

A. OBB based model for object representation

Urban environments are characterised by limited spaces available for navigation and there are little objects movement constraints. In these conditions, geometrical representation of dynamic objects is necessary. Oriented bounding box (OBB) is a way of representing objects geometry with sufficient approximation for the means of navigation.

The OBB based representation is described by two vectors z (1) and σ_z^2 (2). The first one represents the OBB geometry and includes the centre coordinations cx, cy, the orientation angle θ and the size dx, dy. The second vector represents uncertainties on the components of the vector z.

$$z = [cx, cy, \theta, dx, dy]^T \quad (1)$$

$$\sigma_z^2 = [\sigma_{cx}^2, \sigma_{cy}^2, \sigma_\theta^2, \sigma_{dx}^2, \sigma_{dy}^2]^T \quad (2)$$

To construct the OBB based measurement, a specific method is used. The OBB construction method consists of the

978-1-4244-2794-9/09/$25.00 © 2009 IEEE

three following main steps. The first step is to find a contour of the tracked objects using a convex-hull technique [7]. In the second step, a method based on Rotating Calipers (RC) technique [8] is used to construct an OBB, which is best aligned to the object's contour. Finally, the third step concerns the application of the IR uncertainty, which is explained in the section IV-A. The previous steps are described in details in [1].

B. Inter-Rays uncertainty

An important aspect of OBB extraction is the fact that the raw data points representing the extremities of the extracted OBB do not coincide with the real object's extremities (see Figure 1).

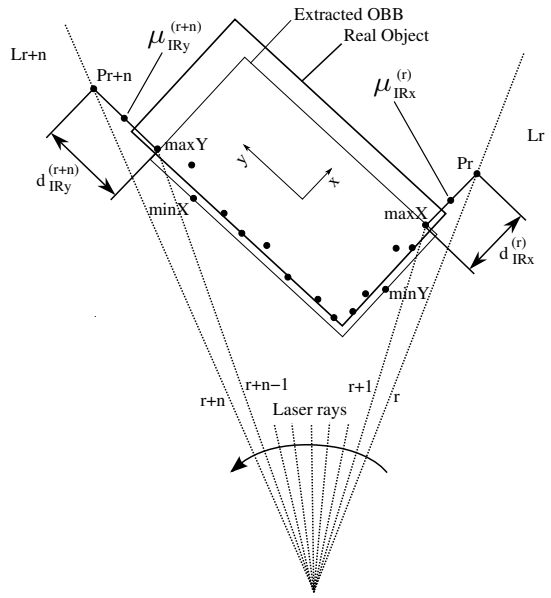


Fig. 1. Inter-Rays uncertainty paradigm.

In the Figure 1, $minX$, $minY$, $maxX$, $maxY$ are respectively the minimum x coordinate, the minimum y coordinate, the maximum x coordinate and the maximum y coordinate of the extracted OBB. The line Lr (respectively $Lr + n$) is crossing the point $maxY$ (respectively $minX$) and is perpendicular to the OBB side to which $maxY$ (respectively $minX$) belongs. The Inter-Rays (IR) real object's extremities position estimation and their variances are added to the OBB's size and OBB's size uncertainty. The real object's extremities are situated between the raw data points delimiting the OBB ($maxY$, $minX$) and the points Pr and $Pr + n$. Pr (respecitvely $Pr + n$) is the intersection point between the ray r (respecitvely $r + n$) with the line Lr (respectively $Lr + n$).

Considering the OBB's local X axis, the real object's extremity position is uniformly distributed with the mean μ_{IRx}, which is equal to the half of the IR line segment length d_{IRx}. The IR line segment is defined by the point $maxY$ and Pr. To fulfil Kalman Filter assumption, the distribution of the real object's extremity position is approximated by a normal distribution with the mean μ_{IRx}, and the variance σ^2_{IRx}, which is set to $\frac{d_{IRx}}{6}$. The Inter-Rays values $z[\mu_{IRx}]$ and $z[\sigma^2_{IRx}]$ are used in each iteration of the tracking algorithm to correct the size of the OBB measurement [1]. The correction equations are expressed as follows:

$$z[dx] = z_{perc}[dx] + z[\mu_{IRx}] \quad (3)$$

$$z[\sigma^2_{dx}] = z_{perc}[\sigma^2_{dx}] + z[\sigma^2_{IRx}] \quad (4)$$

where z_{perc} is the percepred measurement, z is the corrected measurement used for tracking.

The same process is applied for the OBB's local Y axis.

III. Data association

Data association is an important part of multiple-objects tracking. We use laser raw data points clustering with tract-to-cluster correlation to detect the three following possible cases: new object appearance, separate objects tracking, coalescing objects tracking (see Figure 2. Each of the detected cases is treated separately in terms of data association.

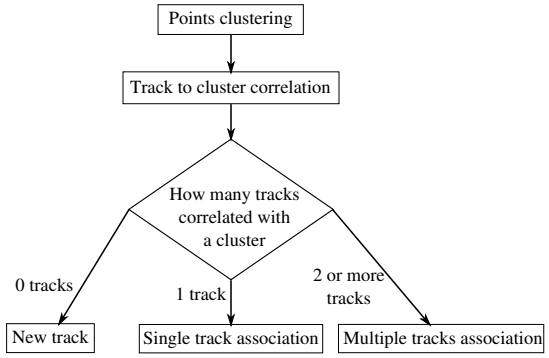


Fig. 2. Data association schema.

The reliability of data association methods depends on the points clustering correctness. In general cases, threshold based clustering gives satisfying results [3], [4], [5]. However, there are situations where, without additional information, it is impossible to cluster raw data points correctly. One of these situations is illustrated in Figure 3. The configuration (a) represents the raw data points corresponding to the rear of a vehicle, seen by the LRF. The configuration (b) could correspond to two possible situations. The first one "vehicle turning" (c) represents the vehicle, which is turning to the right from its former position (a). The second situation "vehicle occlusion" (d) represents two vehicles: the first one (a) and a second vehicle, which is perceived partially by the LRF.

Using only LRF data threshold based clustering, it is impossible to achieve a correct discrimination between different situations. To discard the ambiguities, the authors propose to fuse LRF data and stereovision information. The difference between the proposed approach and standard existing algorithms is that the stereovision information is produced and analyzed only within regions of interest, defined by the projection, onto the stereo images, of couples of consecutive laser points that verify a distance constraint.

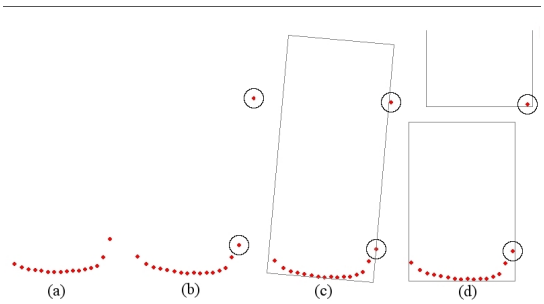


Fig. 3. Laser data clustering ambiguities.

In other terms, for each two points, which are produced by neighbouring laser rays, if the distance between them is greater than a certain threshold, the stereovision analysis is performed to decide if these points belong to a same object. Figure 3 shows a couple of consecutive points (surrounded by circles) for which the stereovision analysis is necessary.

The proposed clustering algorithm is illustrated in Figure 4. The schema shows the clustering process for each input point P_j. Input points are processed consecutively. In the first step, a classical threshold based clustering is performed. If the point P_j is not assigned to any cluster, the existence of a neighbouring point P_i is checked. Neighbouring points are the raw data points produced by neighbouring rays. If the test fails, the point P_j creates a new cluster, otherwise a point gating is achieved. If the point P_j is inside of a track's gate, it is added to the track's cluster, if any or creates a new cluster. If the point P_j is outside of all existing tracks' gates, the stereovision analysis is performed for the points P_j and P_i. Basing on the stereovision analysis, the point P_j creates a new cluster or is added to the cluster of the point P_i.

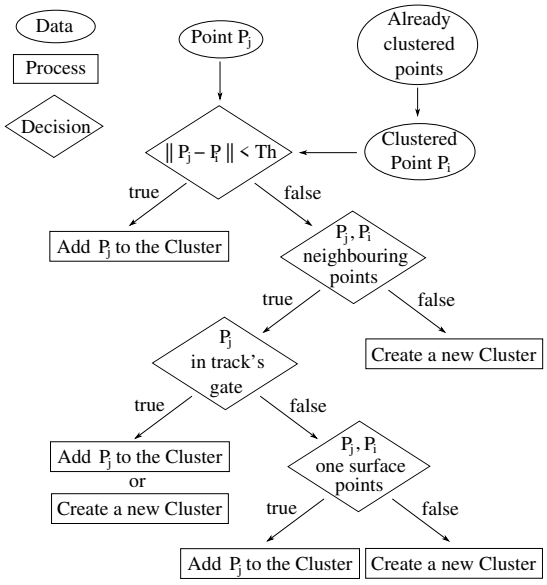


Fig. 4. Clustering schema.

In the following, sections III-A and III-B describe respectively the stereovision method to extract depth information and the fusion algorithm for eliminating ambiguities in the objects clustering procedure.

A. Stereovision method

Providing 3D information of an observed scene, stereovision allows to obtain depth information needed to differentiate different objects when LRF data cannot perform correctly the clustering.

Various correlation methods of grayscales stereoscopic images exist in the literature. The ZSAD (Zero mean Sum of Absolute Differences), which is the most popular correlation method, is chosen because of its less sensitivity to illumination changes [9][10][11]. In this paper, this method is extended to color images. For that, the similarity values are calculated on each component of the RGB color space. The color similarity value corresponds to the average of the three components.

The ZSAD method is based on pattern comparison. For each pixel of an image, a correlation window containing the neighbourhood of the considered pixel is used to compare the similarity. For each correlation window in the left image, the correlation window with the greatest similarity is sought in the right image. The similarity between two correlation windows is expressed as follows:

$$ZSAD(f_l, f_r) = \left\|(f_l - \overline{f_l}) - (f_r - \overline{f_r})\right\| \quad (5)$$

where f_l and f_r denote respectively the vectors containing the pixel values of the correlation window in the left and right images. $\overline{f_l}$ and $\overline{f_r}$ correspond to the average of the pixel values of f_l and f_r vectors respectively.

Figure 5 shows the two clustering ambiguous situations (previously described; see Figure 3), where stereovision information is useful to perform a correct laser data clustering. The first situation concerns a vehicle, which is turning to the right (top-left image). The second one concerns a vehicle partially occluded by another (top-right image). The corresponding disparity maps in grayscales are presented (middle images). In order to visualise better the detected regions, labelled images are extracted (bottom images) from the disparity maps. In these images, each detected region corresponding to a disparity value is represented by a different color.

Recall that for the clustering task, the stereovision analysis is performed only in regions of interest (ROI), defined from the image-projections of couples of consecutive laser points that respect a distance thresholding rule (see section III-B). Considering the example of Figure 5, the ROIs are defined from the image-projections of the couples of the consecutive laser points, surrounded by circles.

Figures 6 shows two surface maps of two horizontal zones corresponding respectively to the two situations of Figure 5.

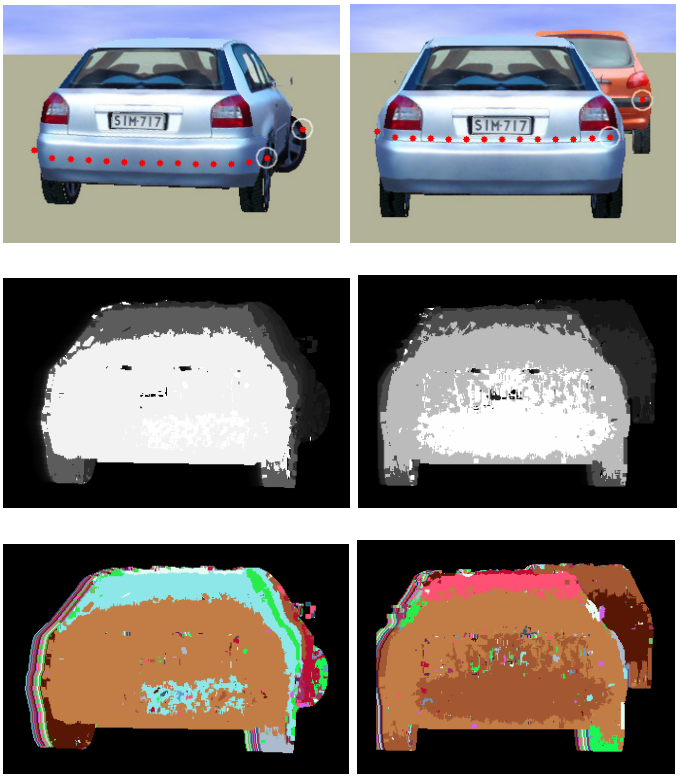


Fig. 5. Disparity maps original images (top), their disparity maps in grayscales (middle) with labelled images b(bottom) for the "vehicle turning" (left) and the "vehicle occlusion" (right) clustering ambiguous situation.

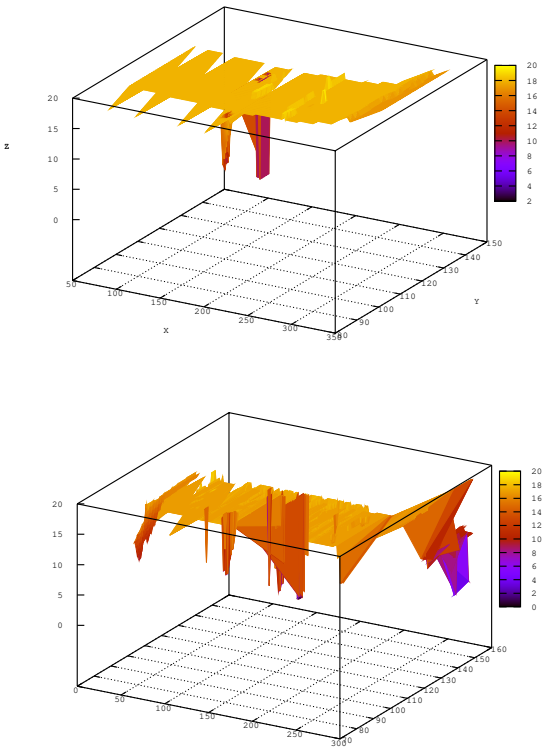


Fig. 6. Surface maps: the "vehicle turning" situation (top); the "vehicle occlusion" situation (left).

The zones are centred on the projections of the laser points onto the images. Each surface map is represented by the disparity values (Z-axis) within the image space (X-axis and Y-axis). Farther is the pixel, smaller is the disparity value (yellow to purple).

In the fist scenario (vehicle which is turning to the right), one can notice that the disparity values are globally constant in the beginning, and then, they decrease gradually between the last two projected laser points, surrounded by circles in Figure 5, and which define the ROI. In the second scenario (vehicle which is partially occluded by another), the same remark can be formulated considering the beginning of the zone, despite some errors near to the registration plate. What is important is that a high gap is visible between the last two projected laser points, surrounded by circles in Figure 5, and which define the ROI.

These statements are more illustrated in Figure 7, which gives the zooms of the disparity surface maps on the neighbourhood of the ROI.

B. *Information fusion*

To decide if two consecutive laser points belong to a same object, the disparity map of the region of interest defined by the projected laser points is analysed. The disparity map analysis consists of detecting discontinuities, between the projected laser points.

To achieve that, a disparity map exploration process is developed in order to check if a discontinuity-less path between the projected laser points exist. If a discontinuity-less path is found, then the two laser points are considered to belong to the same object. Otherwise, the two points belong to different objects.

The search is performed according to a disparity rule: two adjacent pixels p_i and p_j belong to the path if the quantity $\Delta Disp_{ij} = |Disp(p_i) - Disp(p_j)|$ is inferior to a disparity threshold T_{disp}. $Disp(p_i)$ is the p_i pixel disparity value. The disparity threshold T_{disp} is expressed as follows:

$$T_{disp} = \frac{T_r \times T_d \times \Delta Disp_{0N}}{d} \quad (6)$$

where T_r is a weighting coefficient and T_d is a distance threshold. $\Delta Disp_{0N}$ is the difference of the disparity values of the laser points p_0 and p_N, defining the ROI in which the stereovision analysis is performed, and d is the distance between them.

To optimise the disparity map exploration, a specific order is considered, since the disparity maps have always similar topological structure (see Figure 7).

IV. Tracking

A. *Fixed Size assumption*

The idea of the fixed size (FS) assumption is based on the fact that, in general cases, objects' size does not change during

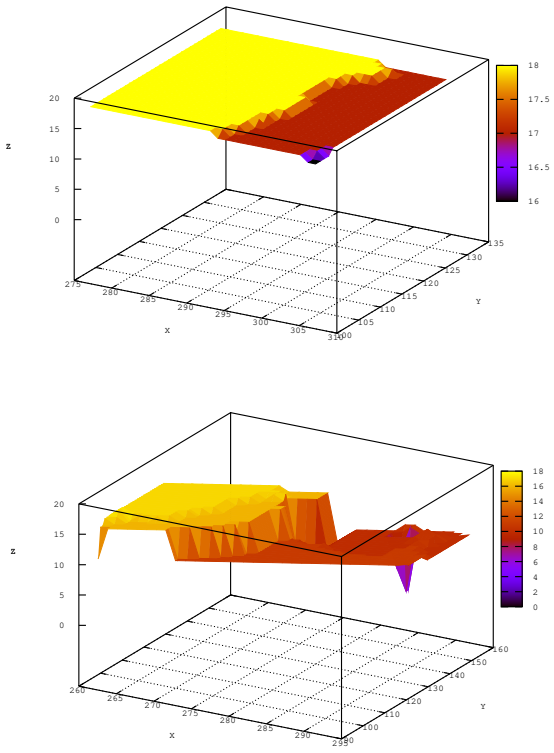


Fig. 7. Zoom of the surface maps around the ROIs.

the tracking. However, due to the LRF's limited resolution and change of the relative distance and orientation of the observed object, measurements of the object's size vary in time. The FS assumption allows to exploit the most precise objects size estimation, memorised during the tracking. The FS algorithm takes place in each iteration of the tracking after the track prediction and measurements extraction.

For the FS algorithm description, we consider the local OBB's X axis. The same process is applied to the local OBB's Y axis.

Having the percepted OBB measurement with the IR line segment length $z_{perc}[d_{IRx}]$, we obtain the corrected IR line segment length $z[d_{IRx}]$ associated with the OBB measurement:

$$z[d_{IRx}] = min(z_{perc}[d_{IRx}], x_{k-1}[d_{IRx}]) \tag{7}$$

where $x_{k-1}[d_{IRx}]$ is the IR line segment length associated with the track at time k-1. The quantity $z[d_{IRx}]$ is then memorised in the track x_k:

$$x_k[d_{IRx}] = z[d_{IRx}] \tag{8}$$

After using the equation (3) and (4) (see section II-B, the next step consists of the measurement's size correction by using the following equation:

$$z[dx] = max(z[dx], x_{\overline{k}}[dx]) \tag{9}$$

where $x_{\overline{k}}[dx]$ is the track predicted size at the time k.

After correcting the percepted measurement's size, the measurement's centre must be appropriately translated. The updating of the centre position is achieved as follows.

Firstly, the visibility factor VF_x is computed for the OBB's local X axis:

$$VF_x = \frac{max(\beta^f_{minX}, \beta^f_{maxX})}{\beta^f_{minX} + \beta^f_{maxX}} \tag{10}$$

where β_{minX} and β_{maxX} correspond respectively to the angles between OBB's sides $minX\,side$ and $maxX\,side$ normals and their radius vectors (see Figure 8). f is a parameter, which is set to 4.

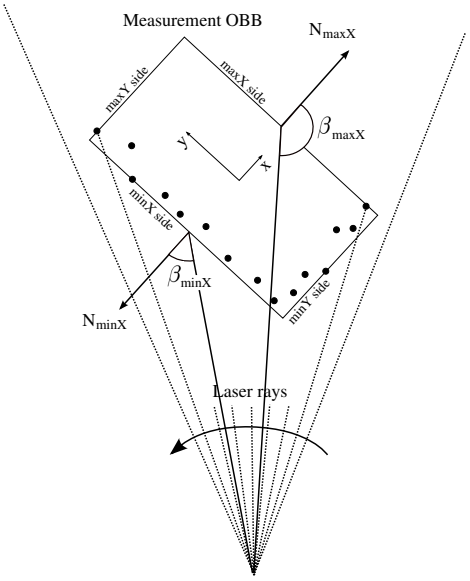


Fig. 8. Visibility factor associated to the OBB's local X axis.

In the second step, the direction factor DF_x associated to the OBB's local X axis is computed using the following equation:

$$DF_x = \begin{cases} 1, & \text{if } \beta_{maxX} > \beta_{minX} \quad (11a) \\ -1, & \text{if } \beta_{maxX} < \beta_{minX} \quad (11b) \end{cases}$$

In the last step, the difference between the percepted size $z_{perc}[dx]$ and the corrected size $z_[dx]$ is calculated:

$$\Delta_{dx} = z[dx] - z_{perc}[dx] \tag{12}$$

Finally, the measurement's centre translation is expressed as follows:

$$\begin{aligned} z[cx] &= z[cx] + \tfrac{1}{2} VF_x \cdot DF_x \cdot \Delta_{dx} \\ z[\sigma^2_{cx}] &= \tfrac{1}{2} z[\sigma^2_{dx}] \end{aligned} \tag{13}$$

B. Extended Kalman Filter

The object's state estimation is done by the means of Extended Kalman Filter (EKF). All values of the track's state vector are expressed in the local ego-vehicle coordinate system. Tracks are represented by the augmented OBB state vector x_k:

$$x_k = [cx, \dot{cx}, cy, \dot{cy}, \theta, \dot{\theta}, dx, dy]^T \quad (14)$$

Since tracking is done from dynamic platform, odometry information is used to increase the tracking accuracy. State change of the ego-vehicle is represented as differences of position $\Delta x, \Delta y$ and angle $\Delta\gamma$ between consecutive instants. Thus, the input to the state transition equation is defined as:

$$u_k = [\Delta x, \Delta y, \Delta\gamma] \quad (15)$$

The Discrete White Noise Acceleration Model (DWNA) [6] is used to describe objects kinematics and process noise. Thus, taking into account the odometry information, the track state transition is modelled as follows [1]:

$$\hat{x}_{\overline{k}|k-1} = A(\Delta x, \Delta y, \Delta\gamma)F\hat{x}_{k-1} + Bu_k + Gv_{k-1} \quad (16)$$

where F is is the standard DWNA transition matrix, B is the odometry-input model, G represents the noise gain matrix, v_{k-1} is the process noise, defined with the Gaussian distribution:

$$v_{k-1} = [\ddot{cx}, \ddot{cy}, \ddot{\theta}, \hat{\sigma}_{\overline{dx}}, \hat{\sigma}_{\overline{dy}}], \qquad v_{k-1} \sim N(0, Q_k) \quad (17)$$

where

$$Q_k = Gv_{k-1}G^T \quad (18)$$

with $\hat{\sigma}_{\overline{dx}}$ and $\hat{\sigma}_{\overline{dy}}$ are the process errors for OBB sizes dx and dy respectively. The prediction covariance matrix is:

$$P_{k|k-1} = \frac{\partial A}{\partial x}(\hat{x}_{k-1})FP_{k-1}\frac{\partial A^T}{\partial x}(\hat{x}_{k-1})F^T + Q_k \quad (19)$$

where P_{k-1} is the estimation covariance matrix.

The observation equation can be written as follows:

$$z_k = H\hat{x}_{\overline{k}|k-1} + w_k \quad (20)$$

where H is the observation model and w_k, which represents the measurement noise, is defined with a Gaussian distribution:

$$w_k \sim N(0, R)R = \sigma_z^2 I_{5,5} \quad (21)$$

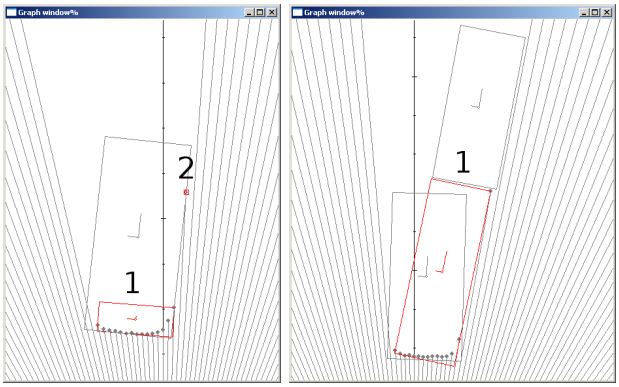
where $I_{5,5}$ is the identity matrix.

V. Experimental results

To evaluate the proposed approach, we use a scenario where the ego vehicle follows another vehicle. During its travel, (see Figure 11), the preceding vehicle avoids a stationary object. At the beginning of the avoidance manoeuvre the clustering ambiguities appear for the two objects seen by the LRF: the preceding vehicle and stationary object.

Figures 9 and 10 present the results of the threshold based clustering for aforementioned clustering ambiguity situations (see sectionIII). The gray rectangle corresponds to the real objects. The gray points represent the LRF raw data points. The red rectangle represents the raw data points cluster. In the case of no-fusion clustering it is difficult to find an unique threshold allowing to achieve a correct clustering for all the situations. In deed, the threshold is to small in the case show in the Figure 9(a) and is to big in case shown in the Figure 10(b).

One can see in Figure 10, that the fusion based clustering algorithm produces correct clusters of all the situations.



(a) (b)

Fig. 9. Data association results of LRF based clustering for the "vehicle turning" situation (a); and for the "vehicle occlusion" one (b).

(a) (b)

Fig. 10. Data association results LRF-stereovision fusion based data clustering for the "vehicle turning" situation (a); and for the "vehicle occlusion" one (b).

Figures 12 and 13 show the estimation errors of the tracked object centre's position, in the X and Y axis respectively. Figures 14 and 15 show the estimation errors of the tracked objects size, in the X and Y axis respectively.

One can see that X coordinate centre's position and objects size estimations are more precise than those associated with the Y coordinate. This is due to the fact, that most of the time, only the X-side of the tracked object is seen. At the beginning of obstacle avoidance manoeuvre (around 600th time instant) the Y coordinate related state estimation becomes more precise thanks to the appearance of the tracked object second side (Y-side). One can see also that the Y coordinate related state

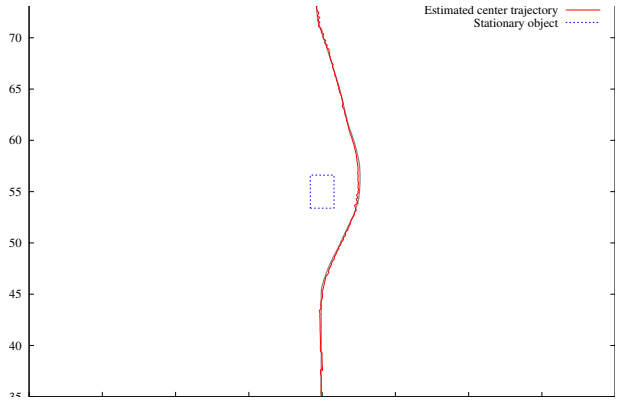


Fig. 11. Trajectory of the tracked object in absolute coordinate system - with LRF-stereovision fusion.

estimation stays almost unchanged in terms of precision, even if again only one side (X-side) of the object is seen. This is due to the FS assumption, which allows to exploit the most precise object's size estimation, memorised during the tracking. The correct object's states estimation and tracking is guaranteed thanks to the correct LRF data clustering by LRF-Stereovision fusion. In deed, without stereovision information and when clustering ambiguity situations appear, the tracking fails (creation of many false tracks).

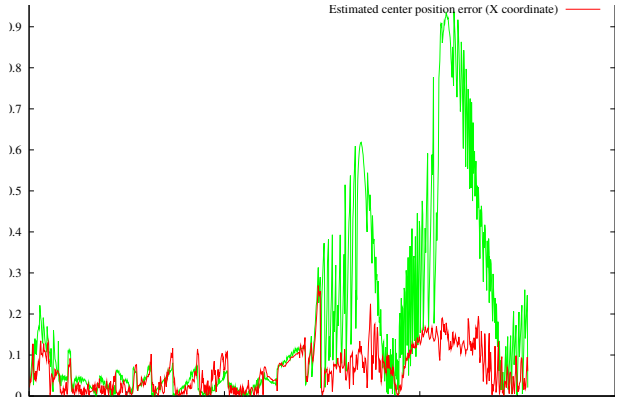


Fig. 12. Tracked object centre's position error - X coordinate.

VI. Conclusion

The problem discussed in this paper concerns objects tracking using laser sensory data and stereovision. The proposed approach uses Oriented Bounding Box (OBB) for representing the tracked objects. The representation model takes into account the Inter-Rays (IR) uncertainty in order to improve the estimation of the extracted OBB. Concerning data association, the authors proposed to integrate stereovision information for extracting laser raw data clusters. The fusion algorithm allows removing the clustering ambiguities pointed out when only LRF data are used. To improve the objects state estimation, the tracking process uses the Fixed Size (FS) assumption, which is introduced to exploit the most precise object's size estimation, memorised during the tracking. The contribution of

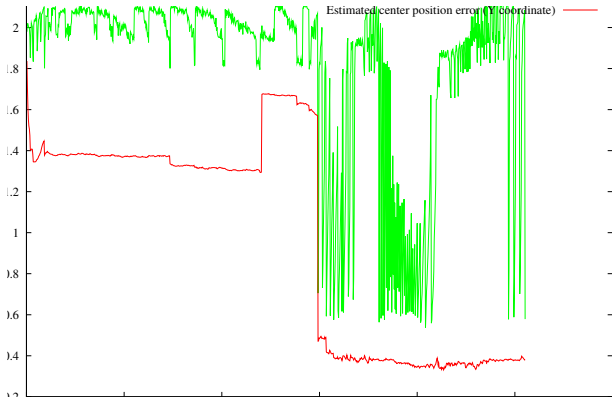


Fig. 13. Tracked object centre's position error - Y coordinate.

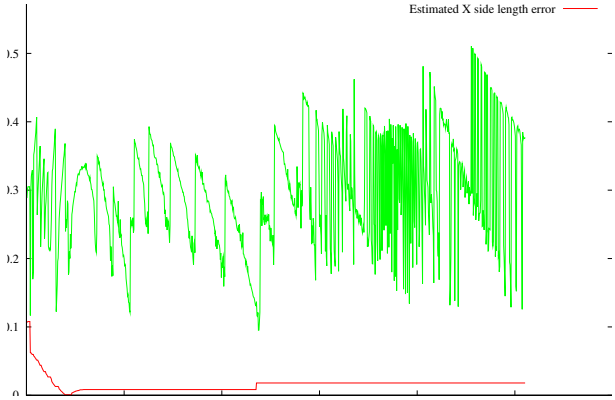


Fig. 14. Tracked object side error - X coordinate.

the IR uncertainty, FS assumption and stereovision information is demonstrated. The presented experimental results show the effectiveness and the reliability of the proposed approach.

References

[1] P. Kmiotek and Y. Ruichek, "Representing and tracking of dynamics objects using oriented bounding box and extended kalman filter." in *Proceedings , 11th International IEEE Conference on Intelligent Transportation Systems*, 2008, pp. 322–328.

[2] A. Petrovskaya and S. Thrun, "Model based vehicle tracking for autonomous driving in urban environments," in *Proceedings of Robotics: Science and Systems Conference 2008*, 2008.

[3] C. Mertz, D. Duggins, J. Gowdy, J. Kozar, R. MacLachlan, A. Steinfeld, A. Suppe, C. Thorpe, and C. Wang, "Collision warning and sensor data processing in urban areas," in *Proceedings of the 5th international conference on ITS telecommunications*, 2005.

[4] K. Dietmayer, J. Sparbert, and D. Streller, "Model based object classification and object tracking in traffic scenes from range images," in *Proc. of 4th IEEE Intelligent Vehicles Symposium*, 2001.

[5] S. Santos, J. Faria, F. Soares, R. Araujo, and U. Nunes, "Tracking of multi-obstacles with laser range data for autonomous vehicles," in *in Proc. of 3rd Nat. Festival of Robotics Scientific Meeting (ROBOTICA'03)*, 2003.

[6] Y. Bar-Shalom, X. Li, and T. Kirubarajan, *Estimation with applications to tracking and navigation*. Wiley New York, 2001.

[7] P. Kmiotek and Y. Ruichek, "Objects oriented bounding box based representation using laser range finder ensory data," in *Proceedings in IEEE International Conference on Vehicular Electronics and Safety*, 2008, pp. 180–184.

[8] G. Toussaint, "Solving geometric problems with the rotating calipers," in *Proc. MELECON, Athens, Greece*, 1983.

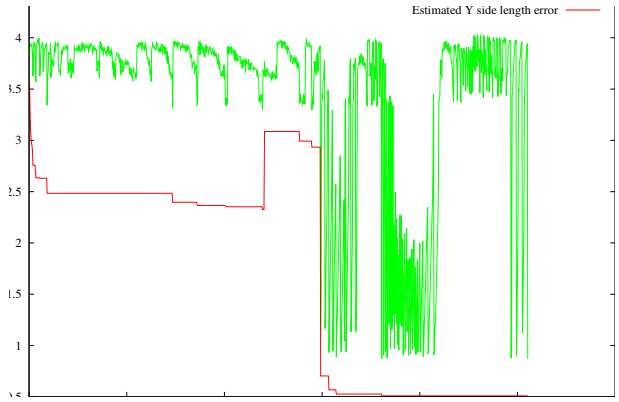


Fig. 15. Tracked object side error - Y coordinate.

[9] L. DiStefano, M. Marchionni, S. Mattocia, and G. Neri, "A fast area-based stereo matching algorithm," in *International Conference on Vision Interface (VI)*, 2002.

[10] H. Hirschmuller, "Improvements in real-time correlation-based stereo-vision." in *Workshop on Stereo and Multi-Baseline Vision*, 2001, pp. 141–148.

[11] C. Kim, K.-M. Lee, B.-T. Choi, and S.-U. Lee, "A dense stereo matching using two-pass dynamic programming with generalized ground control points," in *IEEE Conference of Computer Vision and Pattern Recognition (CVPR)*, vol. 2, 2005, pp. 1075–1082.